This Page Is Inserted by IFW Operations
and is not a part of the Official Record

# BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

## IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

#5 a.a. 11/30/01

PTO/SB/17 (12/99)
Approved for use through 09/30/2000. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

# FEE TRANSMITTAL for FY 2000

*Patent fees are subject to annual revision.*
*Small Entity payments must be supported by a small entity statement, otherwise large entity fees must be paid. See Forms PTO/SB/09-12. See 37 C.F.R. §§ 1.27 and 1.28.*

**TOTAL AMOUNT OF PAYMENT** ($) 950.00

| Complete if Known | |
|---|---|
| Application Number | |
| Filing Date | May 15, 2001 |
| First Named Inventor | Miska HANNUKSELA |
| Examiner Name | |
| Group / Art Unit | |
| Attorney Docket No. | 367.40119X00 |

jc997 U.S. PTO 09/855640 05/15/01

## METHOD OF PAYMENT (check one)

1. ☐ The Commissioner is hereby authorized to charge indicated fees and credit any overpayments to:

Deposit Account Number: 01-2135

Deposit Account Name: Antonelli,Terry,Stout&Kraus,LLP

☒ Charge Any Additional Fee Required Under 37 CFR §§ 1.16 and 1.17

2. ☒ **Payment Enclosed:** ☐ Check ☐ Money Order ☒ Other

## FEE CALCULATION

### 1. BASIC FILING FEE

| Large Entity Fee Code | Large Entity Fee ($) | Small Entity Fee Code | Small Entity Fee ($) | Fee Description | Fee Paid |
|---|---|---|---|---|---|
| 101 | 690 | 201 | 345 | Utility filing fee | 710.00 |
| 106 | 310 | 206 | 155 | Design filing fee | |
| 107 | 480 | 207 | 240 | Plant filing fee | |
| 108 | 690 | 208 | 345 | Reissue filing fee | |
| 114 | 150 | 214 | 75 | Provisional filing fee | |

SUBTOTAL (1) ($) 710.00

### 2. EXTRA CLAIM FEES

| | | Extra Claims | | Fee from below | Fee Paid |
|---|---|---|---|---|---|
| Total Claims | 13 | -20** = 0 | X | 18 | = 0 |
| Independent Claims | 6 | - 3** = 3 | X | 80 | = 240 |
| Multiple Dependent | | | | | = 0 |

**or number previously paid, if greater; For Reissues, see below

| Large Entity Fee Code | Large Entity Fee ($) | Small Entity Fee Code | Small Entity Fee ($) | Fee Description |
|---|---|---|---|---|
| 103 | 18 | 203 | 9 | Claims in excess of 20 |
| 102 | 78 | 202 | 39 | Independent claims in excess of 3 |
| 104 | 260 | 204 | 130 | Multiple dependent claim, if not paid |
| 109 | 78 | 209 | 39 | ** Reissue independent claims over original patent |
| 110 | 18 | 210 | 9 | ** Reissue claims in excess of 20 and over original patent |

SUBTOTAL (2) ($) 240.00

## FEE CALCULATION (continued)

### 3. ADDITIONAL FEES

| Large Entity Fee Code | Large Entity Fee ($) | Small Entity Fee Code | Small Entity Fee ($) | Fee Description | Fee Paid |
|---|---|---|---|---|---|
| 105 | 130 | 205 | 65 | Surcharge - late filing fee or oath | 0.00 |
| 127 | 50 | 227 | 25 | Surcharge - late provisional filing fee or cover sheet. | 0.00 |
| 139 | 130 | 139 | 130 | Non-English specification | 0.00 |
| 147 | 2,520 | 147 | 2,520 | For filing a request for reexamination | 0.00 |
| 112 | 920* | 112 | 920* | Requesting publication of SIR prior to Examiner action | 0.00 |
| 113 | 1,840* | 113 | 1,840* | Requesting publication of SIR after Examiner action | 0.00 |
| 115 | 110 | 215 | 55 | Extension for reply within first month | 0.00 |
| 116 | 380 | 216 | 190 | Extension for reply within second month | 0.00 |
| 117 | 870 | 217 | 435 | Extension for reply within third month | 0.00 |
| 118 | 1,360 | 218 | 680 | Extension for reply within fourth month | 0.00 |
| 128 | 1,850 | 228 | 925 | Extension for reply within fifth month | 0.00 |
| 119 | 300 | 219 | 150 | Notice of Appeal | 0.00 |
| 120 | 300 | 220 | 150 | Filing a brief in support of an appeal | 0.00 |
| 121 | 260 | 221 | 130 | Request for oral hearing | 0.00 |
| 138 | 1,510 | 138 | 1,510 | Petition to institute a public use proceeding | 0.00 |
| 140 | 110 | 240 | 55 | Petition to revive - unavoidable | 0.00 |
| 141 | 1,210 | 241 | 605 | Petition to revive - unintentional | 0.00 |
| 142 | 1,210 | 242 | 605 | Utility issue fee (or reissue) | 0.00 |
| 143 | 430 | 243 | 215 | Design issue fee | 0.00 |
| 144 | 580 | 244 | 290 | Plant issue fee | 0.00 |
| 122 | 130 | 122 | 130 | Petitions to the Commissioner | 0.00 |
| 123 | 50 | 123 | 50 | Petitions related to provisional applications | 0.00 |
| 126 | 240 | 126 | 240 | Submission of Information Disclosure Stmt | 0.00 |
| 581 | 40 | 581 | 40 | Recording each patent assignment per property (times number of properties) | 0.00 |
| 146 | 690 | 246 | 345 | Filing a submission after final rejection (37 CFR § 1.129(a)) | 0.00 |
| 149 | 690 | 249 | 345 | For each additional invention to be examined (37 CFR § 1.129(b)) | 0.00 |
| Other fee (specify) | | | | | 0.00 |
| Other fee (specify) | | | | | 0.00 |

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) ($) 0.00

| SUBMITTED BY | | | | Complete (if applicable) | |
|---|---|---|---|---|---|
| Name (Print/Type) | Carl J. Brundidge | Registration No. (Attorney/Agent) | 29,621 | Telephone | 703-312-6600 |
| Signature | | | | Date | May 15, 2001 |

**WARNING:**

**Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.**

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

THIS PAGE BLANK (USPTO)

367.40119X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

jc997 U.S. PTO
09/855640
05/15/01

Applicant(s): M. HANNUKSELA

Serial No.: Not assigned

Filed: May 15, 2001

Title: VIDEO CODING

Group: Not assigned

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

May 15, 2001

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on United Kingdom Patent Application No.(s) 0011597.2 filed May 15, 2000.

A certified copy of said United Kingdom Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

Carl I. Brundidge
Registration No. 29,621

CIB/amr
Attachment
(703) 312-6600

THIS PAGE BLANK (USPTO)

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

jc997 U.S. PTO
09/855640
05/15/01

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed M Conkie

Dated 2 May 2001

THIS PAGE BLANK (USPTO)

Patents Form 1/77

Patents Act
(Rule 16)

The Patent Office

PATENT OFFICE
A
15 MAY 2000
RECEIVED BY FAX

1/77

# Request for grant of a patent

*(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

PAT 00406 GB 16MAY00 E536859-1 D02716
P01/7700 0.00-0011597.2

2. Patent application number
*(The Patent Office will fill in this part)*

0011597.2

15 MAY 2000

3. Full name, address and postcode of the each applicant *(underline all surnames)*

NOKIA MOBILE PHONES LIMITED
KEILALAHDENTIE 4
02150 ESPOO
FINLAND

Patents ADP number *(if you know it)*

5911995004

If the applicant is a corporate body, give the country/state of its incorporation

FINLAND

4. Title of the invention

VIDEO CODING

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

NOKIA IPR DEPARTMENT
NOKIA HOUSE
SUMMIT AVENUE
FARNBOROUGH
HAMPSHIRE
GU140NG UK

Patents ADP number *(if you know it)*

7577638001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

| Country | Priority application number *(if you know it)* | Date of filing *(day / month / year)* |
|---|---|---|

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

| Number of earlier application | Date of filing *(day / month / year)* |
|---|---|

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

*a) any applicant named in part 3 is not an inventor, or*
*b) there is an inventor who is not named as an applicant, or*
*c) any named applicant is a corporate body.*
*See note (d))*

YES

Patents Form 1/77

**Patents Form 1/77**

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

THE PATENT OFFICE
A
16 MAY 2000
RECEIVED BY FAX

| | |
|---|---|
| Continuation sheets of this form | |
| Description | 26 |
| Claim(s) | 2 |
| Abstract | 1 |
| Drawing(s) | 6 |

10. If you are also filing any of the following, state how many against each item.

| | |
|---|---|
| Priority documents | |
| Translations of priority documents | |
| Statement of inventorship and right to grant of a patent *(Patents Form 7/77)* | |
| Request for preliminary examination and search *(Patents Form 9/77)* | |
| Request for substantive examination *(Patents Form 10/77)* | |
| Any other documents *(please specify)* | |

11. I/We request the grant of a patent on the basis of this application.

Signature [signature] Date 12.06.2000

JULIET HIBBERT

12. Name and daytime telephone number of person to contact in the United Kingdom

Miss J Hibbert 01252 865101

**Warning**

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

**Notes**

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

VIDEO CODING

This invention relates to video coding.

A video sequence consists of a series of still pictures or frames. Video compression methods are based on reducing the redundant and perceptually irrelevant parts of video sequences. The redundancy in video sequences can be categorised into spectral, spatial and temporal redundancy. Spectral redundancy refers to the similarity between the different colour components of the same picture. Spatial redundancy results from the similarity between neighbouring pixels in a picture. Temporal redundancy exists because objects appearing in a previous image are also likely to appear in the current image. Compression can be achieved by taking advantage of this temporal redundancy and predicting the current picture from another picture, termed an anchor or reference picture. Further compression is achieved by generating motion compensation data that describes the motion between the current picture and the previous picture.

However, sufficient compression cannot usually be achieved by only reducing the inherent redundancy of the sequence. Thus, video encoders also try to reduce the quality of those parts of the video sequence which are subjectively less important. In addition, the redundancy of the encoded bit-stream is reduced by means of efficient lossless coding of compression parameters and coefficients. The main technique is to use variable length codes.

Video compression methods typically differentiate between pictures that utilise temporal redundancy reduction and those that do not. Compressed pictures that do not utilise temporal redundancy reduction methods are usually called INTRA or I-frames or I-pictures. Temporally predicted images are usually forwardly predicted from a picture occurring before the current picture and are called INTER or P-frames. In the INTER frame case, the predicted motion-

compensated picture is rarely precise enough and therefore a spatially compressed prediction error frame is associated with each INTER frame. INTER pictures may contain INTRA-coded areas.

Many video compression schemes also use temporally bi-directionally predicted frames, which are commonly referred to as B-pictures or B-frames. B-pictures are inserted between anchor picture pairs of I- and/or P-frames and are predicted from either one or both of these anchor pictures. B-pictures normally yield increased compression as compared with forward-predicted pictures. B-pictures are not used as anchor pictures, i.e., other pictures are not predicted from them. Therefore they can be discarded (intentionally or unintentionally) without impacting the picture quality of future pictures. Whilst B-pictures may improve compression performance as compared with P-pictures, their generation requires greater computational complexity and memory usage, and they introduce additional delays. This may not be a problem for non-real time applications such as video streaming but may cause problems in real-time applications such as video-conferencing.

A compressed video clip typically consists of a sequence of pictures, which can be roughly categorised into temporally independent INTRA pictures and temporally differentially coded INTER pictures. Since the compression efficiency in INTRA pictures is normally lower than in INTER pictures, INTRA pictures are used sparingly, especially in low bit-rate applications.

A video sequence may consist of a number of scenes or shots. The picture contents may be remarkably different from one scene to another, and therefore the first picture of a scene is typically INTRA-coded. There are frequent scene changes in television and film material, whereas scene cuts are relatively rare in video conferencing. In addition, INTRA pictures are typically inserted to stop temporal propagation of transmission errors in a reconstructed video signal and to provide random access points to a video bit-stream.

Compressed video is easily corrupted by transmission errors, mainly for two reasons. Firstly, due to utilisation of temporal predictive differential coding (INTER frames), an error is propagated both spatially and temporally. In practice this means that, once an error occurs, it is easily visible to the human eye for a relatively long time. Especially susceptible are transmissions at low bit-rates where there are only a few INTRA-coded frames, so temporal error propagation is not stopped for some time. Secondly, the use of variable length codes increases the susceptibility to errors. When a bit error alters the codeword, the decoder will lose codeword synchronisation and also decode subsequent error-free codewords (comprising several bits) incorrectly until the next synchronisation (or start) code. A synchronisation code is a bit pattern which cannot be generated from any legal combination of other codewords and such codes are added to the bit stream at intervals to enable re-synchronisation. In addition, errors occur when data is lost during transmission. For example, in video applications using the unreliable UDP transport protocol in IP networks, network elements may discard parts of the encoded video bit-stream.

There are many ways for the receiver to address the corruption introduced in the transmission path. In general, on receipt of a signal, transmission errors are first detected and then corrected or concealed by the receiver. Error correction refers to the process of recovering the erroneous data perfectly as if no errors had been introduced in the first place. Error concealment refers to the process of concealing the effects of transmission errors so that they are hardly visible in the reconstructed video sequence. Typically some amount of redundancy is added by the source or transport coding in order to help error detection, correction and concealment. Error concealment techniques can be roughly classified into three categories: forward error concealment, error concealment by post-processing and interactive error concealment. The term "forward error concealment" refers to those techniques in which the transmitter side adds redundancy to the transmitted data to enhance the error

resilience of the encoded data. Error concealment by post-processing refers to operations at the decoder in response to characteristics of the received signals. These methods estimate the correct representation of erroneously received data. In interactive error concealment, the transmitter and receiver co-operate in order to minimise the effect of transmission errors. These methods heavily utilise feedback information provided by the receiver. Error concealment by post-processing can also be referred to as passive error concealment whereas the other two categories represent forms of active error concealment.

There are numerous known concealment algorithms, a review of which is given by Y. Wang and Q. -F. Zhu in "Error Control and Concealment for Video Communication: A Review", Proceedings of the IEEE, Vol. 86, No. 5, May 1998, pp. 974 – 997 and an article by P. Salama, N. B. Shroff, and E. J. Delp. "Error Concealment in Encoded Video," submitted to IEEE Journal on Selected Areas in Communications.

Current video coding standards define a syntax for a self-sufficient video bit-stream. The most popular standards at the time of writing are ITU-T Recommendation H.263, "Video coding for low bit rate communication", February 1998; ISO/IEC 14496-2, "Generic Coding of Audio-Visual Objects. Part 2: Visual", 1999 (known as MPEG-4); and ITU-T Recommendation H.262 (ISO/IEC 13818-2) (known as MPEG-2). These standards define a hierarchy for bit-streams and correspondingly for image sequences and images.

In H.263, the syntax has a hierarchical structure with four layers: picture, picture segment, macroblock, and block layer. The picture layer data contain parameters affecting the whole picture area and the decoding of the picture data. Most of this data is arranged in a so-called picture header.

The picture segment layer can either be a group of blocks layer or a slice layer. By default, each picture is divided into groups of blocks. A group of

blocks (GOB) typically comprises 16 successive pixel lines. Data for each GOB consists of an optional GOB header followed by data for macroblocks. If the optional slice structured mode is used, each picture is divided into slices instead of GOBs. A slice contains a number of successive macroblocks in scan-order. Data for each slice consists of a slice header followed by data for the macroblocks.

Each GOB or slice is divided into macroblocks. A macroblock relates to 16 x 16 pixels (or 2 x 2 blocks) of luminance and the spatially corresponding 8 x 8 pixels (or block) of chrominance components. A block relates to 8 x 8 pixels of luminance or chrominance.

Block layer data consist of uniformly quantised discrete cosine transform coefficients, which are scanned in zigzag order, processed with a run-length encoder and coded with variable length codes. MPEG-2 and MPEG-4 layer hierarchies resemble the one in H.263.

By default these standards generally use the temporally previous reference picture (I or P) (also known as an anchor picture) as a reference for motion compensation. This piece of information is not transmitted, i.e., the bit-stream does not include information identifying the reference picture. Consequently, decoders have no means to detect if a reference picture is lost. Although many transport coders place video data into packets and associate a sequence number with the packets, these sequence numbers are not related to the video bit-stream. For example, a section of video bit-stream may contain P-picture P1, B-picture B2, P-picture P3, and P-picture P4, captured (and to be displayed) in this order. However, this section would be compressed, transmitted, and decoded in the following order: P1, P3, B2, P4 since B2 requires both P1 and P3 before it can be encoded or decoded. Assuming that there is one picture per packet, that each packet contains a sequence number and that the packet carrying B2 is lost, the receiver can

detect this packet loss from the packet sequence numbers. However, the receiver has no means to detect if it has lost a motion compensation reference picture for P4 or if it has lost a B-picture, in which case it could continue decoding normally.

The decoder therefore usually sends an INTRA request to the transmitter and freezes the picture on the display. However the transmitter may not be about to respond to this request. For instance in a non-real-time video streaming application, the transmitter cannot respond to an INTRA request from a decoder. Therefore the decoder freezes the picture until the next INTRA frame is received. In a real-time application such as video-conferencing, the transmitter may not be able to respond. For instance, in a multi-party conference, the encoder may not be able to respond to individual requests. Again the decoder freezes the picture until an INTRA frame is output by the transmitter.

According to a first aspect of the invention there is provided a method of encoding a video signal representing a sequence of pictures, the method employing both non-temporal prediction and temporal prediction, wherein the method comprises, for each picture that forms a reference picture for the temporal prediction of another picture, associating with each such picture an indicator indicating the temporal order of the reference picture in the encoded video signal relative to the other reference pictures in the encoded video signal.

Thus each reference picture (e.g. I-frames and P-frames) is associated with a sequence number. Preferably the indicator is incremented each time a reference picture is encoded. Most advantageously the indicator is incremented by one each time a reference picture is encoded. Thus the indicator is incremented by one from the previous reference picture.

If multi-layer coding is used, preferably this indicator is incremented by one from the previous reference picture in the same enhancement layer.

Including this indicator means that a decoder is capable of determining whether a reference picture has been lost and to take appropriate action, if available. This is the case even if the transport protocol does not include sequence information about the packets being transmitted or the transmitter uses a varying encapsulation strategy. For example, if the transmitter encapsulates a varying number of GOBs in one packet, receivers cannot reliably know how many GOBs or entire pictures were lost even if they could detect packet losses from packet sequence numbers.

The invention also enables a decoder to differentiate B picture losses from reference picture losses. Consequently, decoders can continue decoding after a B picture loss instead of waiting for the next INTRA picture.

In addition a decoder may continue decoding lower enhancement layers if a reference picture from a higher enhancement layer is lost.

The reference picture order number may be in respect of the whole picture or part of a picture. In the former case, typically the reference picture order number is included in a picture header. In a preferred implementation of the invention, the video signal is encoded according to the H.263 standard and the indicator is included in the Supplemental Enhancement Information. Where the RPON is in respect of part of a picture the reference picture order number may be included in the picture segment headers or macroblock headers of the encoded picture.

In accordance with a second aspect of the invention there is provided a method of decoding an encoded video signal representing a sequence of pictures, the method comprising receiving an encoded video signal, decoding each received picture, examining for each picture to be decoded that forms a

reference picture for another picture an indicator representing the temporal order of a reference frame and, when the indicator does not follow consecutively from an immediately preceding decoded reference frame, detecting a lost reference frame.

Preferably the decoder sends a request to a transmitter to encode a picture in a non-temporal manner when the indicator does not follow consecutively from an immediately preceding decoded reference frame.

In accordance with a third aspect of the invention there is provided an encoded video signal wherein for each picture that forms a reference picture for the temporal prediction of another picture, an indicator is associated with each such reference picture indicating the temporal order of the reference pictures in the encoded video stream.

In accordance with a fourth aspect of the invention there is provided a video encoder comprising an input for receiving a video signal representing a sequence of pictures and for generating encoded pictures, said encoder being arranged to employ both non-temporal prediction and temporal prediction, wherein the encoder is arranged, for each picture that forms a reference picture for the temporal prediction of another picture, to associate with each reference picture an indicator indicating the temporal order of the reference picture in the encoded video signal relative to other reference pictures in the encoded video signal.

Preferably the indicator is incremented each time a reference picture is encoded.

In accordance with a fifth aspect of the invention there is provided a video decoder comprising an input for receiving an encoded video signal representing a sequence of pictures, a decoder for decoding each received picture, the decoder being arranged to examine for each picture to be

decoded that forms a reference picture for another picture an indicator representing the temporal order of a reference frame and, when the indicator does not follow consecutively from an immediately preceding decoded reference frame, to detect a lost reference frame.

The invention also relates to a portable radio communications device including an encoder and/or a decoder as described.

The invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows a multimedia mobile communications system;

Figure 2 shows an example of the multimedia components of a multimedia terminal;

Figure 3 shows an example of a video codec;

Figure 4 illustrates the operation of a first implementation of a video encoder according to a first embodiment of the invention;

Figure 5 illustrates the operation of a second implementation of a video encoder according to a first embodiment of the invention;

Figure 6 shows the syntax of a bit stream as known according to H.263;

Figure 7 shows a first example of a bit stream output by an encoder according to the invention;

Figure 8 shows a second example of a bit stream output by an encoder according to the invention;

Figure 9 shows a third example of a bit stream output by an encoder according to the invention;

Figure 10 illustrates enhancement layers used in video coding for SNR scalablity; and

Figure 11 illustrates enhancement layers used in video coding for spatial scalablity.

Figure 1 shows a typical multimedia mobile communications system. A first multimedia mobile terminal 1 communicates with a second multimedia mobile

terminal 2 via a radio link 3 to a mobile communications network 4. Control data is sent between the two terminals 1,2 as well as the multimedia data.

Figure 2 shows the typical multimedia components of a terminal 1. The terminal comprises a video codec 10, an audio codec 20, a data protocol manager 30, a control manager 40, a multiplexer/demultiplexer 50 and a modem 60 (if the required). The video codec 10 receives signals for coding from a video capture device of the terminal (not shown) (e.g. a camera) and receives signals for decoding from a remote terminal 2 for display by the terminal 1 on a display 70. The audio codec 20 receives signals for coding from the microphone (not shown) of the terminal 1 and receive signals for decoding from a remote terminal 2 for reproduction by a speaker (not shown) of the terminal 1.

The control manager 40 controls the operation of the video codec 10, the audio codec 20 and the data protocol manager 30. However, since the invention is concerned with the operation of the video codec 10, no further discussion of the audio codec 20 and protocol manager 30 will be provided.

Figure 3 shows an example of a video codec 10 according to the invention. The video codec comprises an encoder part 100 and a decoder part 200. The encoder part 100 comprises an input 101 for receiving a video signal from a camera or video source (not shown) of the terminal 1. A switch 102 switches the encoder between an INTRA-mode of coding and an INTER-mode.

In INTRA-mode, the video signal from the input 101 is transformed into DCT co-efficients by a DCT transformer 103. The DCT coefficients are then passed to a quantiser 104 that quantises the coefficients. Both the switch 102 and the quantiser 104 are controlled by an encoding control manager 105 of the video codec, which also receives feedback control from the receiving terminal 2 by means of the control manager 40.

In INTER mode, the switch 102 is operated to accept from a subtractor 106 the difference between the signal from the input 101 and a previous picture which is stored in a picture store 107. The difference data output from the subtractor 106 represents the prediction error between the current picture and the previous picture stored in the picture store 107. The data in the picture store 107 is generated by passing the data output by the quantiser through an inverse quantiser 108 and applying an inverse DCT transform 109 to the inverse-quantised data. The resulting data is added to the contents of the picture store 107 by adder 110. A motion estimator 111 may generate motion compensation data from the data in the picture store 107 in a conventional manner.

The encoding control manager 105 decides whether to apply INTRA or INTER coding or whether to code the frame at all on the basis of either the output of the subtractor 106 or in response to feedback control data received a receiving decoder. When not responding to feedback control data, the encoder typically encodes a frame as an INTRA-frame either only at the start of coding (all other frames being P-frames), or at regular periods e.g. every 5s, or when the output of the subtractor exceeds a threshold i.e. when the current picture and that stored in the picture store 107 are too dissimilar. The encoder may also be programmed to encode frames in a particular regular sequence e.g. I B B P B B P B B P B B P B B I B B P etc. In addition the encoding control manager may decide not to code a received frame at all. This happens when the similarity between the current frame and the reference frame is so high that the encoder decides not to encode the current frame. The encoding control manager operates the switch accordingly.

The video codec outputs the quantised DCT coefficients 112a, the quantising index 112b (i.e. the details of the quantiser used), an INTRA/INTER flag 112c to indicate the mode of coding performed (I or P/B), a transmit flag 112d to indicate the number of the frame being coded and the motion vectors 112e for

the picture being coded. These are multiplexed together by the multiplexer 50 together with other multimedia signals.

The decoder part 200 of the video codec 10 comprises an inverse quantiser 120, an inverse DCT transformer 121, a motion compensator 122, a picture store 123 and a controller 124. The controller 124 receives video codec control signals demultiplexed from the encoded multimedia stream by the demultiplexer 50. In practice the controller 105 of the encoder and the controller 124 of the decoder may be the same processor.

The operation of an encoder according to the invention will now be described. The video codec 10 receives a video signal to be encoded. The encoder 100 of the video codec encodes the video signal by performing DCT transformation, quantisation and motion compensation. The decoded video data is then output to the multiplexer 50. The multiplexer 50 multiplexes the video data from the video codec 10 and control data from the control manager 40 (as well as other signals as appropriate) into a multimedia signal. The terminal 1 outputs this multimedia signal to the receiving terminal 2 via the modem 60 (if required).

In a first embodiment of the invention, each time the encoder encodes a frame which may form the reference frame for a subsequent frame, the encoding control manager 105 associates with the frame a so-called Reference Picture Order Number (RPON). For example, a RPON is associated with every I or P frame of a video signal but not with a B-frame. The RPON value is incremented each time a successive reference picture is encoded, preferably by 1.

The encoding control manager 105 outputs the RPON codeword on output 112f which indicates the Reference Picture Order Number associated with the encoded frame. This is multiplexed into the video bitstream by a multiplexer.

Figure 4 illustrates the operation of the encoder. In this embodiment, the encoder is arranged to output an I-frame when the similarity between the frame being coded and the reference frame is less than a first threshold i.e. when the output from the subtractor 106 is greater than a first threshold. Otherwise the encoder outputs P-frames. The first line of Figure 4 represents the frames of data received from a capture input device and input to the video encoder on input 101. The second line of Figure 4 represents those frames of the input signal that the encoder decides to encode and the coding mode used to encode each frame. As mentioned above some, the encoding control manager may decide that a frame is not to be coded: this is exemplified in Figure 4 by the fact that frames 2, 3 and 11 are not coded.

Frame 1 is coded in INTRA-mode; frame 4 is encoded as a P-frame with reference to frame 1; frame 5 is encoded as a P-frame with reference to frame 4; frame 6 is encoded as a P-frame with reference to frame 5; frame 7 is encoded as a P-frame with reference to frame 6; frame 8 is encoded as an I-frame; frame 9 is encoded as a P-frame with reference to frame 8; frame 10 is encoded as a P-frame with reference to frame 9; frame 12 is encoded as a P-frame with reference to frame 10.

In this embodiment all (but the last) of the encoded frames act as the reference frame for a later frame. Thus a RPON is associated with all of the frames to be coded, as shown in the third line of Figure 4. As can be seen, the RPON is incremented by 1 each time.

The fourth line of Figure 4 shows the Temporal Reference (TR) of the encoded frame. This is a field included in H.263 and the value of TR is formed by incrementing its value in the temporally previous reference picture header by one plus the number of skipped or non-reference pictures since the previously transmitted one. Thus in the example shown in Figure 4 the TR shown for each frame is the same as the original number in the original signal input to 102.

Figure 5 shows a second embodiment of an encoder according to the invention. In this embodiment, the encoder is arranged to code the frames according to the regular sequence I B B P B B P B B P B B P B B I B B P. The first line of Figure 5 shows the input frames and the second line shows the coded frames and their coding mode, I, P or B.

The frames are received from a video capture device in the order 1,2,3,4,5,6 etc. and are displayed in this order i.e. the decoded frames are displayed in the order I1,B2,B3,P4,B5,B6,P7 etc. However the video bit stream is compressed, transmitted and decoded in the following order I1, P4, B2, B3, P7, B5, B6 etc. This is because each B-frame requires preceding and succeeding reference frames before they can be encoded/decoded i.e. frame B2 requires frame I1 and P4 to be encoded/decoded before frame B2 can be encoded/decoded.

As explained previously, B-frames are inserted between anchor picture pairs of I- and/or P-frames and are predicted from either one or both of these anchor pictures. Thus in the illustration given in Figure 5, Frame 1 is coded in INTRA-mode; frame 4 is encoded as a B-frame with reference to frame 1 and/or 6; frame 5 is encoded as a B-frame with reference to frame 1 and/or 6; frame 6 is encoded as a P-frame with reference to frame 1; frame 7 is encoded as a B-frame with reference to frame 6 and/or 9; frame 8 is encoded as an B-frame with reference to frame 6 and/or 9; frame 9 is encoded as a P-frame with reference to frame 6; frame 10 is encoded as a B-frame with reference to frame 9 and/or 13 (not shown); frame 12 is encoded as a B-frame with reference to frame 9 and/or 13 and so on.

In this embodiment each I-frame and P-frame of the encoded sequence acts as a reference frame for another frame. However a B-frame does not act as a reference picture for any other frame. Thus a RPON is associated with all of the I-frames and P-frames, as shown in the third line of Figure 5. As can be

seen, the RPON is incremented by 1 each time. Thus frame 1 (an I-frame) has a RPON of 1, frame 4 (a P-frame) has a RPON of 2 and frame 9 (a P-frame) has a RPON of 3.

The fourth line of Figure 5 shows the Temporal Reference (TR) of the encoded frame. As in the example shown in Figure 4, the TR shown for each frame is the same as the order of occurrence in the original signal input to 10.

Considering the terminal 1 as receiving coded video data from terminal 2, the operation of the video codec 10 will now be described with reference to its decoding role. The terminal 1 receives a multimedia signal from the transmitting terminal 2. The demultiplexer 50 demultiplexes the multimedia signal and passes the video data to the video codec 10 and the control data to the control manager 40. The decoder 200 of the video codec decodes the encoded video data by inverse quantising, inverse DCT transforming and motion compensating the data. The controller 124 of the decoder checks the integrity of the received data and, if an error is detected, attempts to conceal the error in a manner to be described below. The decoded, corrected and concealed video data is then output for reproduction on a display 70 of the receiving terminal 1.

Errors in video data may occur at the picture level, the GOB level or the macroblock level. Error checking may be carried out at any or all of these levels.

Considering first the signal as shown in Figure 4, when a decoder according to the invention receives this signal each frame of the signal is decoded in a conventional manner and then displayed on a display means. The decoded frame may be error corrected and error concealed in a conventional manner. Each time a frame is decoded, the decoder examines the TR field to determine when the frame is to be displayed. If the TRs are not consecutive (e.g. the decoder receives a frame with TR=1 and then a frame with TR=4)

the decoder holds the frame 1 on the display for 3 times the usual frame period, as is conventional. The decoder also examines the RPON of the received frames. In the case shown in Figure 4 the decoder receives frame 1 and sees that this frame has a RPON=1; the decoder then receives a frame with TR=4 and RPON=2. The decoder compares the RPON of the currently received frame with the RPON of the previously received frame and calculates the difference between the RPON values. In this case the difference is 1 and the decoder therefore knows that no reference pictures have been lost between the current frame and the previous decoded reference frame. The decoder therefore continues to decode the signal in a conventional manner.

Let us now assume that the decoder is unable to reconstruct frame 5 (this could be due to the data being greatly corrupted or being lost altogether) and the next frame received and decoded by the decoder is frame 6. The decoder compares the RPON of the currently received frame (frame 6) with the RPON of the previously received and decoded reference frame (frame 4) and calculates the difference between the RPON values. In this case the difference is 2 and the decoder therefore knows that a reference picture has been lost between transmission of the current frame and that of the previous frame. If the decoder has the facility to send control feedback data to the transmitting video encoder the decoder can send a request to the transmitting video encoder to encode a frame as an INTRA-frame and so stop the temporal error propagation that would result from frame 6 being decoded with reference to frame 4.

Considering now the signal as shown in Figure 5, when a decoder according to the invention receives this signal each frame of the signal is decoded in a conventional manner and then displayed on a display means. The decoded frame may be error corrected and error concealed in a conventional manner. Each time a frame is decoded, the decoder examines the TR field to

determine when the frame is to be displayed. The decoder also examines the RPON of the received frames.

In the case shown in Figure 5 the decoder receives frame 1 and sees that this frame has a RPON=1. The decoder decodes this frame in a conventional INTRA-mode manner. The next frame received by the decoder is then frame 6, with TR=6 and RPON=2. The decoder compares the RPON of the currently received frame (frame 6) with the RPON of the previously received and decoded reference frame (frame 1) and calculates the difference between the RPON. In this case the difference is 1 and the decoder therefore knows that no reference pictures have been lost between transmission of the current frame and that of the previous decoded reference frame. The decoder then decodes frame 6 with reference to frame 1.

The decoder then receives a frame with TR=4 and no RPON. In this case the decoder makes no further use of the RPON and decodes frame 4 with reference to decoded frames 1 and 6.

Let us now assume that the decoder is unable to reconstruct frame 5 (this could be due to the data being greatly corrupted or being lost altogether). The fact that B-frame 5 has been lost is of no consequence to the decoder as the B-frame does not form a reference picture for any other frame and thus its loss will not introduce any temporal error propagation.

The next frame to be received is frame 9. However, let us now assume that the decoder is unable to reconstruct frame 9, which is a P-frame (this could be due to the data being greatly corrupted or being lost altogether). The decoder may therefore be unable to decode successfully any of frames 7, 8, 10 or 12 since these may all be predicted, in part at least, with reference to frame 9. Typically, in this situation, the decoder will freeze the displayed picture.

The next frame received and decoded by the decoder is frame 13. The decoder compares the RPON of the currently received reference frame (frame 13) with the RPON of the previously received and decoded reference frame (frame 6) and calculates the difference between the RPON. In this case the difference is 2 and the decoder therefore knows that a reference picture has been lost between the current frame and the previous decoded reference frame. If the decoder has the facility to send control feedback data to the transmitting video encoder the decoder can send a request to the transmitting video encoder to encode a frame as an INTRA-frame and so stop the temporal error propagation that would result from frame 13 being decoded with reference to frame 6.

How the reference picture order number may be included in the encoded signal will now be addressed with reference to the H.263 video coding standard.

Figure 6 shows the syntax of a bit stream as known according to H.263. The following implementations describe the GOB format but it will be clear to a skilled person that the invention may also be implemented in the slice format.

As mentioned already, the bit stream has four layers: the picture layer, picture segment layer, macroblock layer and block layer. The picture layer comprises a picture header followed by data for the Group of Blocks, eventually followed by any optional end-of-sequence code and stuffing bits.

The prior art H.263 bit stream is formatted as shown in Figure 6. A descriptor for each part is given below:

PSC The picture start code (PSC) indicates the start of the picture

TR The Temporal Reference (TR) is formed by incrementing its value in the temporally previous reference picture header by one plus the number of skipped or non-referenced pictures since the previously transmitted one

| | |
|---|---|
| PTYPE | Amongst other things, PTYPE includes details of the picture coding type i.e. INTRA or INTER |
| PQUANT | A codeword that indicates the quantiser to be used for the picture until updated by any subsequent quantiser information |
| CPM | A codeword that signals the use of optional continuous presence multipoint and video multiplex (CPM) mode |
| PSBI | Picture Sub-Bit stream Indicator – only present if CPM is set |
| $TR_B$ | Present if the frame is a bi-directionally predicted frame (known as a PB-frame) |
| DBQUANT | Present if a bi-directional frame |
| PEI | This relates to extra insertion information and is set to "1" to indicate the presence of the following optional data fields PSUPP and PEI. PSUPP and PEI are together known as Supplemental Enhancement Information, which is further defined in Annex L of H263. |
| GOBS | Is the data for the group of blocks for the current picture |
| ESTF | A stuffing codeword provided to attain byte alignment before EOS |
| EOS | A codeword indicating the end of the data sequence for the picture |
| PSTUF | A stuffing codeword to allow for byte alignment of the next picture start code PSC |

The structure as shown in Figure 4 does not include the optional PLUSTYPE data field. PSBI is only present if indicated by CPM. TRB and DBQUANT are only present if PTYPE indicates use of a so-called PB frame mode (unless the PLUSTYPE filed is present and the used of DBQUANT is indicated therein). These issues are addressed in more detail in the H.263 specification.

The following paragraphs outline possible implementations of the bit-stream output by an encoder according to the first aspect of the invention.

The reference picture order number may be incorporated into a H.263 bit stream as follows. Figure 7 shows an example of a bit stream output by an encoder according to the first implementation of the invention. As shown in Figure 7, the bit stream includes a further codeword RPON which is a codeword indicating the reference picture order number. This is inserted by an encoder according to the invention, as described above.

Alternatively, the reference picture order number may be included in the Supplemental Enhancement Information PSUPP (see Annex L of H.263 and Figure 4). The supplemental information may be present in the bit stream even though the decoder may not be capable of providing the enhanced capability to use it, or even to properly interpret it. Simply discarding the supplemental information is allowable by decoders unless a requirement to provide the requested capability has been negotiated by the transmitter and receiver.

If PEI is set to "1", then 9 bits follow consisting of 8 bits of data (PSUPP) and then another PEI bit to indicate if a further 9 bits follow and so on.

The PSUPP data consists of a 4-bit function type indication FTYPE, followed by a 4-bit parameter data size specification DSIZE followed by DSIZE octets of function parameter data, optionally followed by another FTYPE and so on. It is known to use this PSUPP codeword to signal various situations such as: to indicate a full-picture or partial-picture freeze or freeze-release request with or without resizing; to tag particular pictures or sequences of pictures within the video stream for external use; or to convey chroma key information for video compositing.

To implement the invention using the Supplemental Enhancement Information, a further FTYPE is defined as Reference Picture Order Number.

Figure 8 illustrates the example where a parameter RPON is included in the SEI of the picture header. The FTYPE is defined as RPON. The DSIZE specifies the size of the parameter and the following octet is the parameter data i.e. the value of RPON. From this value a receiving decoder can determine whether a reference picture has been lost.

Alternatively, the information may be contained in the additional Supplemental Enhancement Information as specified in a "Draft of new Annex W: Additional Supplementary Enhancement Information Specification" P. Ning and S. Wenger, ITU-T Study Group 16 Question 15 Document Q15-I-58, November 1999.

In this draft proposal, FTYPE 14 is defined as "Picture Message". When this FTYPE is set, the picture message function indicates the presence of one or more octets representing message data. The first octet of the message data is a message header with the structure shown in Figure 9 i.e. CONT, EBIT and MTYPE. DSIZE is equal to the number of octets in the message data corresponding to a picture message function, including the first octet message header.

The continuation field CONT, if equal to 1, indicates that the message data associated with the picture message is part of the same logical message as the message data associated with the next picture message function. The End Bit Position field EBIT specifies the number of least significant bits that shall be ignored in the last message octet. Further details of these fields can be found in the draft of Annex W referred to above.

The field MTYPE indicates the type of message. Various types of message are suggested in the draft of Annex W. According to the invention one type e.g. MTYPE 9 is defined as RPON. The value of RPON is defined in the octet following the message header.

In a specific example, this message contains one data byte, i.e., DSIZE is 2, CONT is 0, and EBIT is 0.

The Reference Picture Order Number is incremented by one from the corresponding number of the previous coded reference picture. The least significant 8-bits of the result of the incrementation is placed in the data byte associated with this message.

The above description has made reference to encoded video streams in which bi-directionally predicted pictures (B-pictures) are encoded. As mentioned earlier, B-pictures are never used as reference pictures. Since they can be discarded without impacting the picture quality of future pictures, they provide temporal scalability. Scalability allows for the decoding of a compressed video sequence at more than one quality level. In other words, a scalable multimedia clip can be compressed so that it can be streamed over channels with different data rates and still be decoded and played back in real-time.

Thus the video stream may be decoded in different ways by differing decoders. For instance, a decoder can decide only to decode the I- and P-pictures of a signal, if this is the maximum rate of decoding that the decoder can attain. However if a decoder has the capacity, it can also decode the B-pictures and hence increase the picture display rate. Thus the perceived picture quality of the displayed picture will be enhanced over a decoder that only decodes the I- and P-pictures.

Scalable multimedia is typically ordered so that there are hierarchical layers of data. A base layer contains a basic representation of the multimedia clip whereas enhancement layers contain refinement data on top of underlying layers. Consequently, the enhancement layers improve the quality of the clip.

Scalability is a desirable property for heterogeneous and error prone environments. This property is desirable in order to counter limitations such as constraints on bit rate, display resolution, network throughput, and decoder complexity.

Scalability can be used to improve error resilience in a transport system where layered coding is combined with transport prioritisation. The term transport prioritisation here refers to various mechanisms to provide different qualities of service in transport, including unequal error protection, to provide different channels having different error/loss rates. Depending on their nature, data are assigned differently. For example, the base layer may be delivered through a channel with a high degree of error protection, and the enhancement layers may be transmitted through more error-prone channels.

Generally, scalable multimedia coding suffers from a worse compression efficiency than non-scalable coding. In other words, a multimedia clip encoded as a scalable multimedia clip with enhancement layers requires greater bandwidth than if it had been coded as a non-scalable single-layer clip with equal quality. However, exceptions to this general rule exist, for example the temporally scalable B-frames in video compression.

The invention may be applied to other scalable video compression systems. For instance, in H.263 Annex O, two other forms of scalability are defined: signal-to-noise (SNR) scalability and spatial scalability.

Spatial scalability and SNR scalability are closely related, the only difference being the increased spatial resolution provided by spatial scalability. An example of SNR scalable pictures is shown in Figure 10. SNR scalability implies the creation of multi-rate bit streams. It allows for the recovery of coding errors, or differences between an original picture and its reconstruction. This is achieved by using a finer quantiser to encode the

difference picture in an enhancement layer. This additional information increases the SNR of the overall reproduced picture.

Spatial scalability allows for the creation of multi-resolution bit streams to meet varying display requirements and/or constraints. A spatially scalable structure is illustrated in Figure 11. It is essentially the same as in SNR scalability except that a spatial enhancement layer attempts to recover the coding loss between an up-sampled version of the reconstructed reference layer picture and a higher resolution version of the original picture. For example, if the reference layer has a quarter common intermediate format (QCIF) resolution, and the enhancement layer has a common intermediate format (CIF) resolution, the reference layer picture must be scaled accordingly such that the enhancement layer picture can be predicted from it. The QCIF standard allows the resolution to be increased by a factor of two in the vertical direction only, horizontal direction only, or both the vertical and horizontal directions for a single enhancement layer. There can be multiple enhancement layers, each increasing the picture resolution over that of the previous layer. The interpolation filters used to up-sample the reference layer picture are explicitly defined in the H.263 standard. Aside from the up-sampling process from the reference to the enhancement layer, the processing and syntax of a spatially scaled picture are identical to those of an SNR scaled picture.

In either SNR or spatial scalability, the enhancement layer pictures are referred to as EI- or EP-pictures. If the enhancement layer picture is upwardly predicted from a picture in the reference layer, then the enhancement layer picture is referred to as an Enhancement-I (EI) picture. In this type of scalability, the reference layer means the layer "below" the current enhancement layer. In some cases, when reference layer pictures are poorly predicted, over-coding of static parts of the picture can occur in the enhancement layer, causing an unnecessarily excessive bit rate. To avoid this problem, forward prediction is permitted in the enhancement layer. A picture

that can be predicted in the forward direction from a previous enhancement layer picture or, alternatively, upwardly predicted from the reference layer picture is referred to as an Enhancement-P (EP) picture. Note that computing the average of the upwardly and forwardly predicted pictures can provide bi-directional prediction for EP-pictures. For both EI- and EP-pictures, upward prediction from the reference layer picture implies that no motion vectors are required. In the case of forward prediction for EP-pictures, motion vectors are required.

According to the invention, if the encoder is capable of multi-layer coding (for example as discussed in Annex O of H.263) the reference pictures of each layer are given consecutive Reference Picture Order Numbers. These may be associated with the enhancement layer number (ELNUM) of the current picture. The Reference Picture Order Number is incremented by one from the corresponding number of the previous coded reference picture in the same enhancement layer.

As shown in Figures 10 and 11, the pictures of the enhancement layer may be predicted from a preceding picture of the enhancement layer and/or from the equivalent I- or P-picture of the base layer. The enhancement layer may not be predicted from a B-picture in the reference layer.

If adjacent pictures in the same enhancement layer have the same temporal reference, and if Annex N or Annex U of H.263 is in use, the decoder preferably regards this occurrence as an indication that redundant copies have been sent of approximately the same pictured scene content, and all of these pictures then share the same RPON.

A decoder according to the invention, on receipt of a multi-layer signal as described above, attempts to decode the signal in a conventional manner. In each layer, each time a reference picture is decoded, the decoder examines the RPON of the decoded picture. If the decoder determines that a reference

picture has been lost from an enhancement layer, the decoder ceases to display pictures from the enhancement layer until an EI-picture is received. The decoder continues to decode the base layer as described earlier.

The invention is not intended to be limited to the video coding protocols discussed above: these are intended to be merely exemplary. The invention is applicable to any video coding protocol in which temporal prediction may be used. The addition of the information as discussed above allows a receiving decoder to determine that a reference picture has been lost and to take appropriate action.

CLAIMS

1. A method of encoding a video signal representing a sequence of pictures, the method employing both non-temporal prediction and temporal prediction, wherein the method comprises, for each picture that forms a reference picture for the temporal prediction of another picture, associating with each such picture an indicator indicating the temporal order of the reference picture in the encoded video signal relative to the other reference pictures in the encoded video signal.

2. A method according to claim 1 further comprising incrementing the indicator each time a reference picture is encoded.

3. A method according to claim 2 further comprising incrementing the indicator by one each time a reference picture is encoded.

4. A method according to any preceding claim wherein the indicator is included in a picture header.

5. A method according to claim 3 wherein the video signal is encoded according to the H.263 standard and the indicator is included in the Supplemental Enhancement Information.

6. A method of decoding an encoded video signal representing a sequence of pictures, the method comprising receiving an encoded video signal, decoding each received picture, examining for each picture to be decoded that forms a reference picture for another picture an indicator representing the temporal order of a reference frame and, when the indicator does not follow consecutively from an immediately preceding decoded reference frame, detecting a lost reference frame.

7. A method according to claim 6 further comprising sending a request to a transmitter to encode a picture in a non-temporal manner when the indicator does not follow consecutively from an immediately preceding decoded reference frame.

8. An encoded video signal wherein for each picture that forms a reference picture for the temporal prediction of another picture, an indicator is associated with each such reference picture indicating the temporal order of the reference pictures in the encoded video stream.

9. A video encoder comprising an input for receiving a video signal representing a sequence of pictures and for generating encoded pictures, said encoder being arranged to employ both non-temporal prediction and temporal prediction, wherein the encoder is arranged, for each picture that forms a reference picture for the temporal prediction of another picture, to associate with each reference picture an indicator indicating the temporal order of the reference picture in the encoded video signal relative to other reference pictures in the encoded video signal.

10. A video encoder according to claim 9 further comprising means for incrementing the indicator each time a reference picture is encoded.

11. A video decoder comprising an input for receiving an encoded video signal representing a sequence of pictures, a decoder for decoding each received picture, the decoder being arranged to examine for each picture to be decoded that forms a reference picture for another picture an indicator representing the temporal order of a reference frame and, when the indicator does not follow consecutively from an immediately preceding decoded reference frame, to detect a lost reference frame.

12. A portable radio communications device including an encoder according to claim 9 or 10 and/or a decoder according to claim 11.

ABSTRACT

VIDEO CODING

A method of encoding a video signal representing a sequence of pictures, the method employing both non-temporal prediction and temporal prediction, wherein the method comprises, for each picture that forms a reference picture for the temporal prediction of another picture, associating with each such picture an indicator indicating the temporal order of the reference picture in the encoded video signal relative to the other reference pictures in the encoded video signal.

Fig 5

THIS PAGE BLANK (USPTO)

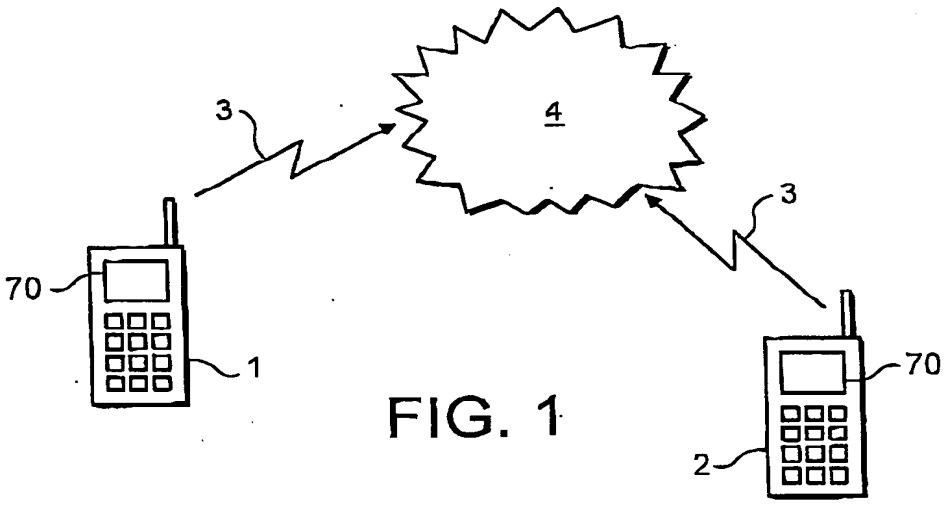



FIG. 1

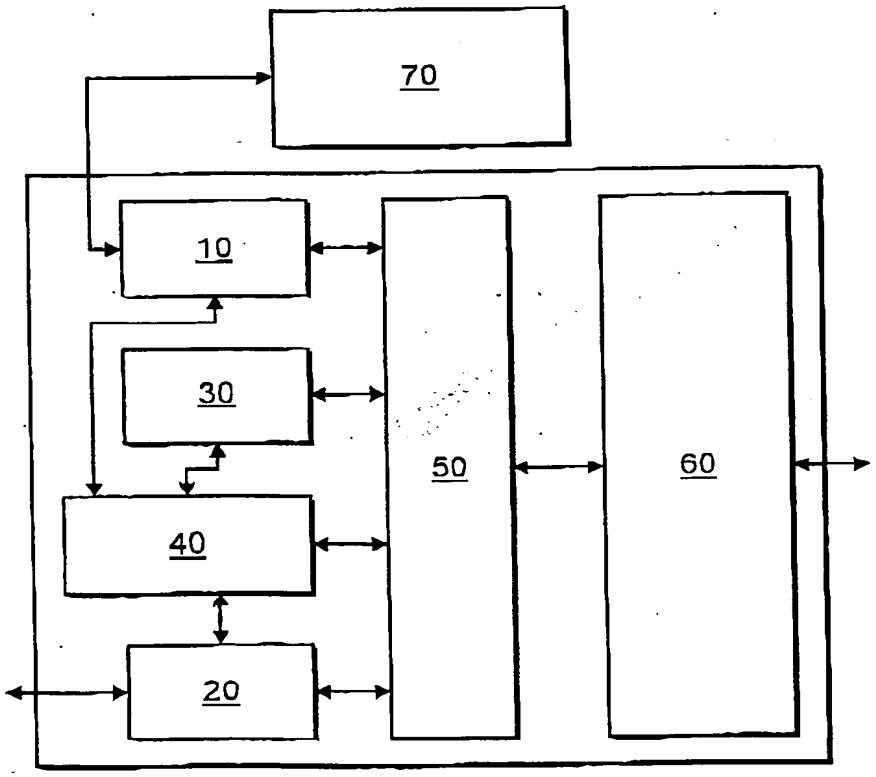



FIG. 2

THIS PAGE BLANK (USPTO)

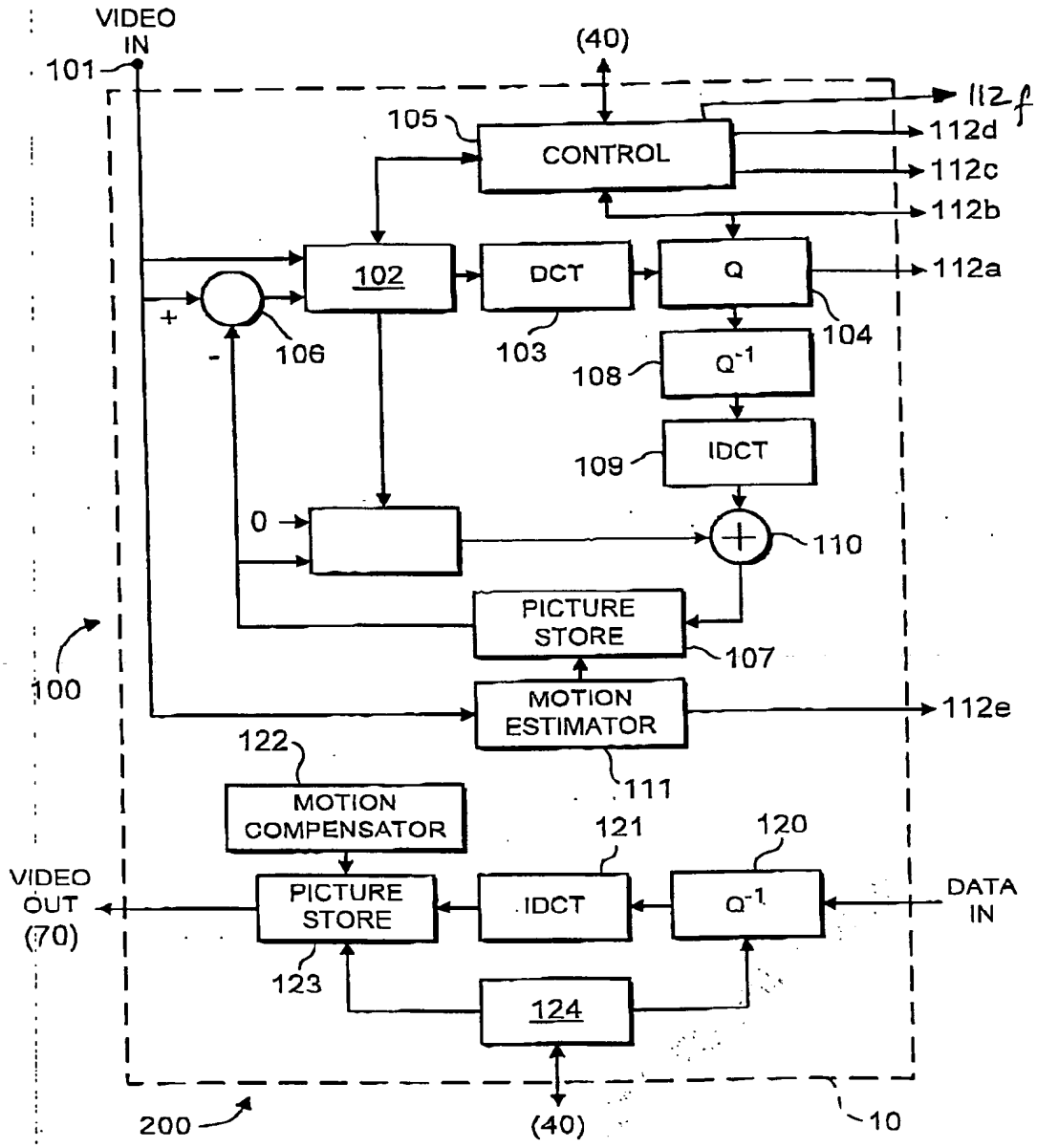

VIDEO IN
101
(40)
112f
105
CONTROL
112d
112c
112b
102
DCT
Q
112a
+
-
106
103
104
108
Q-1
109
IDCT
0
110
PICTURE STORE
107
100
MOTION ESTIMATOR
112e
122
111
MOTION COMPENSATOR
121
120
VIDEO OUT (70)
PICTURE STORE
IDCT
Q-1
DATA IN
123
124
200
(40)
10

FIG. 3

THIS PAGE BLANK (USPTO)

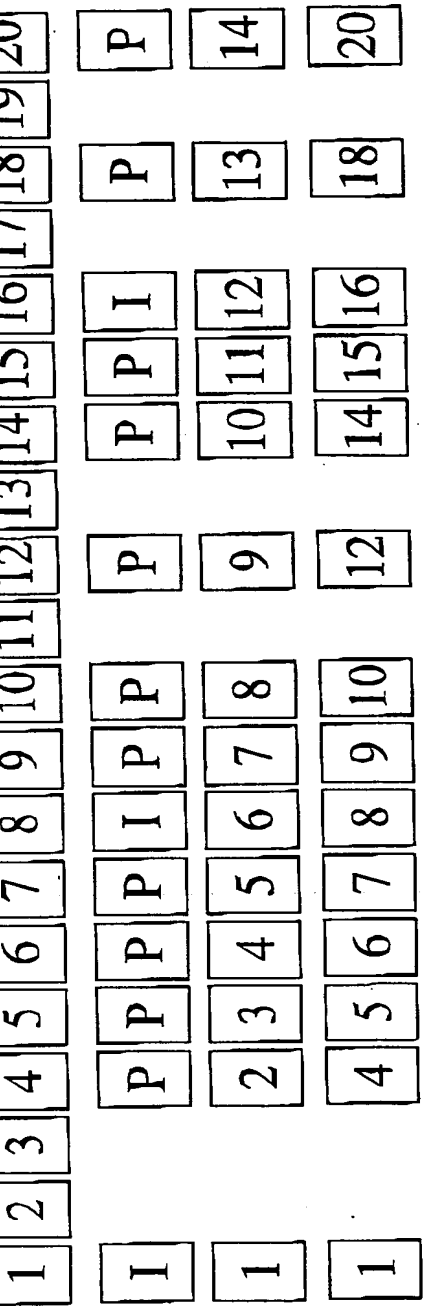

Frame No
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Encoded Frames
I P P P P I P P P P P P P P P P P P P P
RPON
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
TR
1 4 5 6 7 8 9 10 12 14 15 16 18 20

Fig. 4

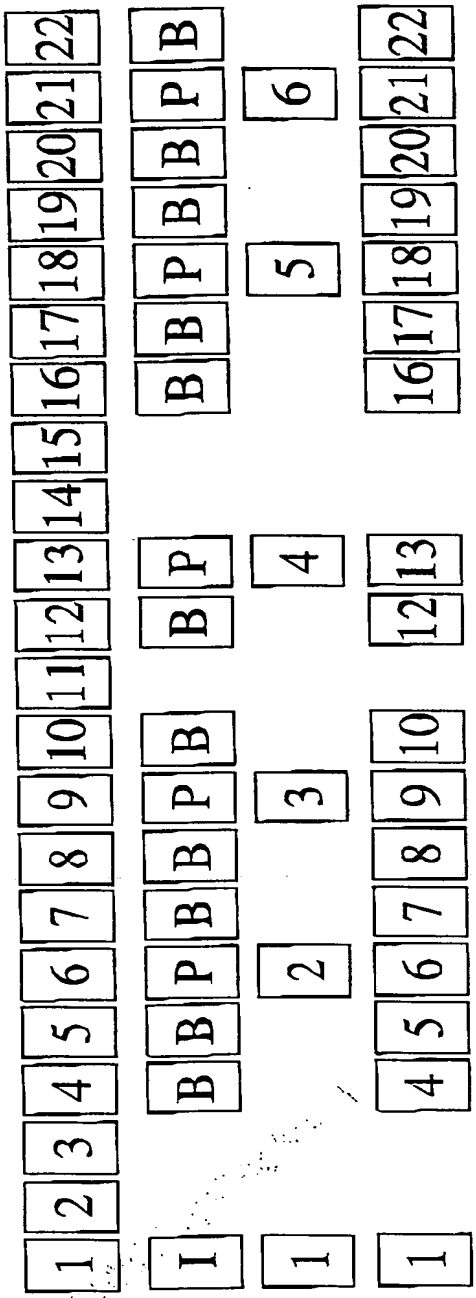

Frame No
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Encoded Frames
I B B P B B P B P B B P B B P B B P B
RPON
1 2 3 4 5 6
TR
1 4 5 6 7 8 9 10 12 13 16 17 18 19 20 21 22

Fig. 5

THIS PAGE BLANK (USPTO)

PSC
TR
PTYPE
PQUANT
CBM
PSBI
$TR_B$
DQUANT
PEI
PSUPP
PEI
GOBs
ESTUF
EOS
PSTUF

Fig 6

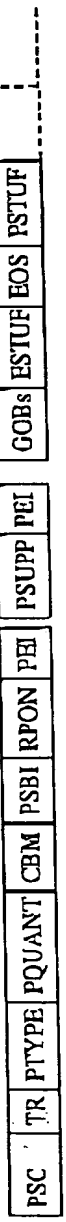

PSC
TR
PTYPE
PQUANT
CBM
PSBI
RPON
PEI
PSUPP
PEI
GOBs
ESTUF
EOS
PSTUF

Fig 7

THIS PAGE BLANK (USPTO)

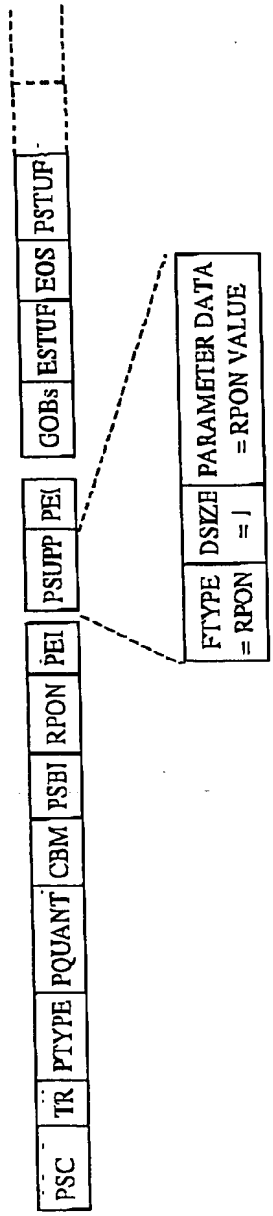

PSC
TR
PTYPE
PQUANT
CBM
PSBI
RPON
PEI
PSUPP
PEI
GOBs
ESTUF
EOS
PSTUF
FTYPE = RPON
DSIZE = 1
PARAMETER DATA = RPON VALUE

Fig 8

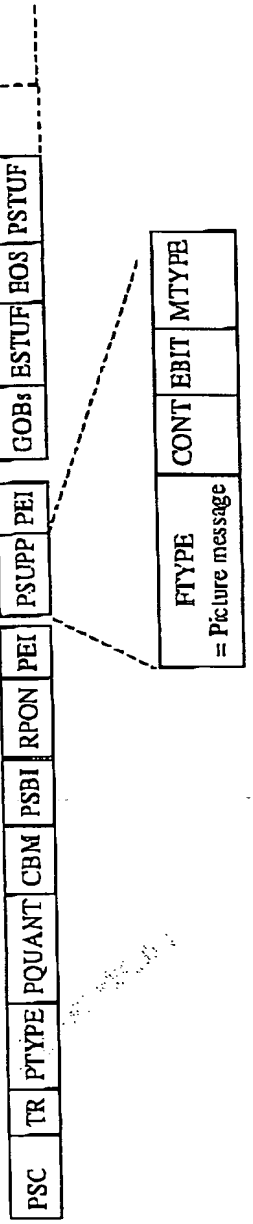

PSC
TR
PTYPE
PQUANT
CBM
PSBI
RPON
PEI
PSUPP
PEI
GOBs
ESTUF
EOS
PSTUF
FTYPE = Picture message
CONT
EBIT
MTYPE

Fig. 9

THIS PAGE BLANK (USPTO)

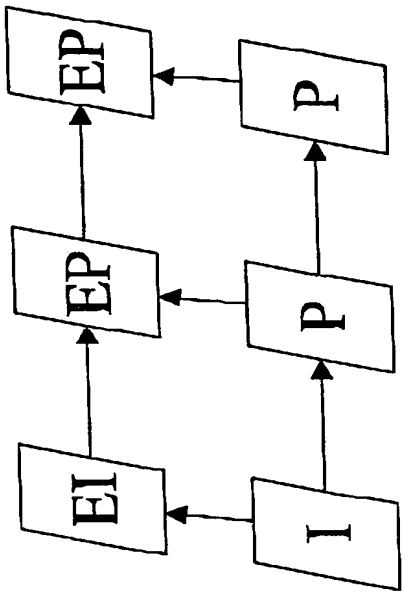



Fig. 10

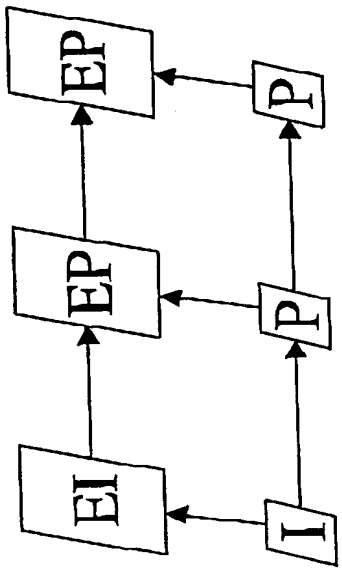
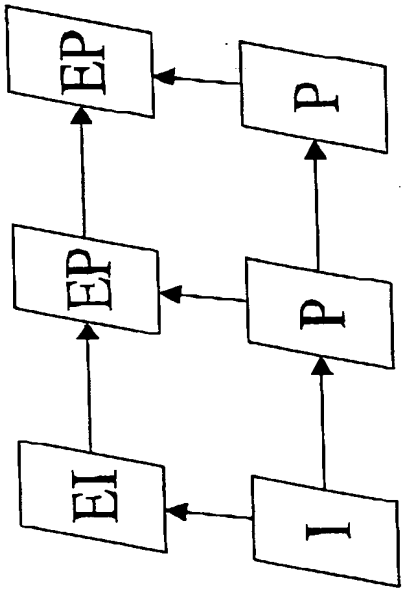
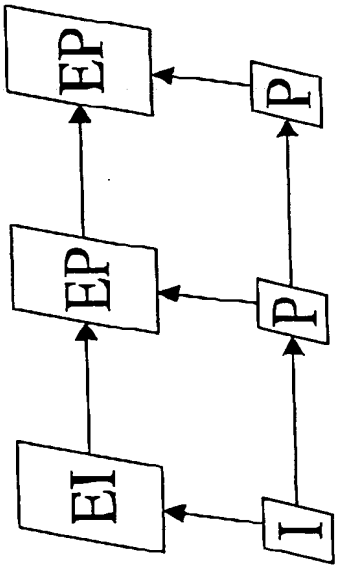



Fig. 11



THIS PAGE BLANK (USPTO)